DOCUMENT RESUME

ED 464 226 CE 083 166

AUTHOR Waugh, C. Keith
TITLE Isolating the Effects of Training Using Simple Regression Analysis: An Example of the Procedure.
PUB DATE 2001-12-00
NOTE 23p.; Paper presented at the Annual Meeting of the Association for Career and Technical Education (75th, New Orleans, LA, December 13-16, 2001).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Education; *Banking; Case Studies; *Cost Effectiveness; *Job Training; *Outcomes of Education; Predictive Measurement; *Predictor Variables; Program Effectiveness; *Regression (Statistics); Salesmanship; Staff Development; Vocational Education
IDENTIFIERS Performance Indicators; *Return on Investment

ABSTRACT

This paper provides a case example of simple regression analysis, a forecasting procedure used to isolate the effects of training from an identified extraneous variable. This case example focuses on results of a three-day sales training program to improve bank loan officers' knowledge, skill-level, and attitude regarding solicitation and sale of loans to new business accounts. It outlines 10 procedural steps taken to identify the most appropriate performance indicator of training success; isolate the effect the three-day training program had on the performance indicator from the most potent extraneous variable; and calculate the return on investment (ROI) of the training program. The 10 procedural steps are the following: (1) identify performance indicator and extraneous factor to be measured; (2) calculate correlation between the performance indicator and the extraneous variable; (3) calculate the slope of the regression equation; (4) calculate intercept [a] of the regression equation; (5) calculate the predicted value of the performance indicator [y] given an extraneous variable value [x]; (6) calculate standard error of estimate [See]; (7) calculate standard error of forecast [Sef]; (8) calculate a confidence interval around the predicted value; (9) calculate monetary benefits of the training program; and (10) calculate ROI. The conclusion is the monetary benefits of the training produced a ROI of 21.19 percent; for every dollar invested in the training program, $0.21 was returned beyond cost of the training. (YLB)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 464 226

ISOLATING THE EFFECTS OF TRAINING USING SIMPLE REGRESSION ANALYSIS: AN EXAMPLE OF THE PROCEDURE

C. Keith Waugh

Southern Illinois University at Carbondale

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

C.K. Waugh

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at the

2001 Association for Career and Technical Education Conference

December, 2001

CE083166

BEST COPY AVAILABLE

Introduction

Many organizations are now demanding evidence from their HRD departments that the money spent on training programs delivered to their employees are producing a return on investment (ROI). However, isolating the effects of training from other non-training factors [extraneous variables] is a challenge for evaluators. Phillips (1997) writes:

> When a significant increase in performance is noted after a major training program has been conducted, the two events appear to be linked. While the change in performance may be linked to the training program, other non-training factors usually have contributed to the improvement as well. The challenge is to develop one or more specific strategies to isolate the effects of training. The cause and effect relationship between training and performance can be very confusing and difficult to prove, but this can be accomplished with an acceptable degree of accuracy. (p. 165)

Various approaches to isolate the effects of training from extraneous variables have been offered. One of these is to compare the after-training performance of the training participants [experimental group] with a control group. In an evaluation design, a control group would consist of employees who do not receive training (or who receive an alternative form of training) but are essentially equal to the experimental group on all other salient characteristics. Since any extraneous variable will have equal influence on the control and experimental group's performance, any performance differences between the groups is attributed to training.

In practice, however, incorporating control groups into an evaluation design is inherently problematic. First, forming a control group of employees essentially equal to the experimental group on the salient characteristics of work experience, pre-training knowledge and skill levels, attitude, motivation, etc. can be a daunting task. One approach to achieve relative equality

between groups is to randomly assign employees to either the control group or experimental group. However, work schedules and responsibilities of employees may prevent the ability for random assignment. Second, from a business perspective, it is often impractical to intentionally withhold training from those who need it for the purpose of forming a control group.

Phillips (1997) offers a second approach that involves asking managers, supervisors, subject-matter experts, customers, subordinates and training participants to estimate the impact that the training has had on improved performance, and to indicate their level of confidence in their estimate. While this approach can be more easily implemented than using a control group evaluation design, these estimates are highly subjective. Thus, the accuracy of the estimates is a concern.

A more viable approach is to utilize business forecasting techniques to isolate the effects of training from extraneous variables that may also influence performance indicators. Essentially, this approach provides an objective, analytical estimate of the level of performance that would have been achieved by the training participants had training not occurred, thus partitioning the influence of an extraneous variable on a performance indicator from the effects of training. Because this approach utilizes objective measures of performance instead of subjective estimates by stakeholders, the results produced from the analysis are more credible. Also, the problems associated with a control group evaluation design are avoided since this approach does not require the formation and inclusion of a control group.

What follows is a case example of a forecasting procedure used to isolate the effects of training from an identified extraneous variable. Numerous forecasting methods are available to evaluators. However, discussing each of the various methods is beyond the scope of this paper. Interested readers should refer to Makridakis and Wheelwright (1989) to learn about other

forecasting methods. In this case example, the forecasting procedure is limited to one forecasting method, simple regression analysis.

Case Example

Background

The interest earned from the outstanding principal of personal loans sold by its loan officers represents a significant source of revenue for a small bank in the mid-west. However, compared to other banks of similar size in the region, the cumulative dollar value of loans sold per month by this bank has been less than competitive. In response, the bank's HRD department conducted a thorough needs analysis and concluded that the bank's 15 loan officers lacked sufficient knowledge, skill, and attitude needed to attract new customers and sell personal loans. Therefore, the HRD department developed and delivered during the first 3 days of July 2001 a 3 day sales training program to improve the loan officers' knowledge, skill-level, and attitude regarding the solicitation and sale of loans to new business accounts. The cost of the needs analysis, development, delivery and evaluation of the program totaled $15,000.

What follows are the procedural steps taken to: 1) identify the most appropriate performance indicator of training success; 2) isolate the effect that the 3 day training program had on the performance indicator from the most potent extraneous variable, and; 3) calculate the return on investment (ROI) of the training program.

Step 1: Identify the Performance Indicator and Extraneous Factor to be Measured

The first step in this evaluation process is to identify the performance indicator and the extraneous variable that has the greatest influence on performance. The performance indicator is the desired outcome from the training. In this case example, the desired outcome is an increase in interest revenue on the outstanding principal of the loans sold. Therefore, the performance indicator is the dollar amount of loans sold among the training participants.

Given the type of training program delivered, the performance indicator is often obvious. However, identifying an extraneous variable that influences the performance indicator is sometimes more difficult. The evaluator should seek assistance from company personnel to help identify the most influential extraneous variable. Gathering as much input as possible from a variety of knowledgeable personnel can be very useful (e.g., department managers, supervisors, accountants, sales, etc.).

During a meeting with the loan officers, the HRD staff and members of management, it was decided that the interest rate established by the bank's board of directors is an extraneous variable that has a profound effect of the dollar amount of loans sold to customers. Specifically, as interest rates rise, it is believed that the amount of loans sold fall.

Step 2: Calculate the Correlation Between the Performance Indicator and the Extraneous Variable

As noted in step 1, the performance indicator is the dollar amount of loans sold and the extraneous variable identified as having the most influence on the amount of loans sold is the interest rate for the loan. To make certain that there is, indeed, a relationship between the two, a correlation coefficient should be calculated.

In Table 1 is historical data showing the total dollar amount of loans sold by the loan officers (y) and the respective interest rate (x) for each of 30 months prior to the delivery of the training in July of 2001.

Statistical software is certainly useful when wanting to calculate a correlation coefficient. However, it is not difficult to calculate a correlation using a simple calculator. Following Table 1 is the formula used for the calculation of a correlation coefficient.

Table 1

Historical data of Total Dollar Amount of Loans Sold by 15 Loan Officers (y) and Interest Rates (x) for Each Month Prior to Training

Month	Interest Rate (in percent) (x)	Total Loans Sold (in thousands) (y)
Jan.99	7.75	22
Feb.99	7.75	20
Mar.99	7.75	21
Apr.99	7.75	20
May99	7.75	22
Jun.99	7.75	20
Jul.99	7.75	21
Aug.99	8	18
Sep.99	8.06	16
Oct.99	8.25	15
Nov.99	8.25	16
Dec.99	8.37	13
Jan.00	8.5	11
Feb.00	8.5	10
Mar.00	8.73	10
Apr.00	8.83	9
May00	9	8
Jun.00	9.24	7
Jul.00	9.5	5
Aug.00	9.5	6
Sep.00	9.5	4
Oct.00	9.5	5
Nov.00	9.5	4
Dec.00	9.5	3
Jan.01	9.05	6
Feb.01	8.5	8
Mar.01	8.32	9
Apr.01	7.8	14
May01	7.24	16
Jun.01	7	20

Correlation Coefficient

$$r = \frac{N(\sum xy) - (\sum x)(\sum y)}{\sqrt{[N(\sum x^2) - (\sum x)^2][N(\sum y^2) - (\sum y)^2]}}$$

To arrive at the numbers to place into the formula above, several calculations need to be made first. When statistical software or a calculator with statistical functions is not available, creating a table can make the task much easier. Table 2 provides the calculations from the data from Table 1. Please note that throughout the remainder of this paper, these same numbers will be used repeatedly. Therefore, it is important that care be used when calculating these numbers.

$$r = \frac{30(3069.69) - (252.89)(379)}{\sqrt{[30(2147.7749) - (252.89)^2][30(5955) - (379)^2]}} = -.92$$

The sum totals of each column from Table 2 have been placed into the correlation coefficient equation shown above, with the exception of *N* which is the number of data points collected for analysis. In this example, *N* is equal to 30 (i.e., 30 months of data).

As shown, the correlation coefficient between the dollar amount of loans sold by the 15 loan officers (y) and interest rates (x) is -.92. This should be interpreted as a strong negative correlation. Thus, as interest rates rise the amount of loans sold falls. Because the correlation is

so strong (only .08 from a perfect negative relationship of -1), it appears that interest rates do have considerable influence on the performance indicator. Therefore, we can feel comfortable continuing the analysis using interest rates as the extraneous variable. Had we found that there is a weak or no relationship between the performance indicator and interest rates, then another extraneous variable should be tested.

Step 3: Calculate the slope of the regression equation

Feeling comfortable that the most influential extraneous variable on the performance indicator has been identified, we now begin a series of calculations that will ultimately lead to a prediction of what the dollar amount of loans sold by the 15 loan officers would have been in July 2001 (the first month after training) had training not occurred, given the loan interest rate for that month. The first calculation to be made in the series is the slope (b) of the regression equation. Below is the formula used:

Slope (b)

$$b = \frac{N(\sum xy) - (\sum x)(\sum y)}{N(\sum x^2) - (\sum x)^2}$$

$$b = \frac{30(3069.69) - (252.89)(379)}{30(2147.7749) - (252.89)^2} = -7.823817$$

Again, take note that the sum totals of the appropriate columns from Table 2 have been placed into the equation. The result is a slope of -7.823817. This number represents the degree

in which the performance indicator (y) changes given a unit change in the extraneous variable (x).

Table 2

Calculations and Sums of Raw Data

Month	Interest Rate (in percent) (x)	x^2	Total Loans Sold (in thousands) (y)	y^2	xy	$(x - \bar{x})^2$
Jan.99	7.75	60.06	22	484	170.5	.46
Feb.99	7.75	60.06	20	400	155	.46
Mar.99	7.75	60.06	21	441	162.75	.46
Apr.99	7.75	60.06	20	400	155	.46
May99	7.75	60.06	22	484	170.5	.46
Jun.99	7.75	60.06	20	400	155	.46
Jul.99	7.75	60.06	21	441	162.75	.46
Aug.99	8	64	18	324	144	.18
Sep.99	8.06	64.96	16	256	128.96	.14
Oct.99	8.25	68.06	15	225	123.75	.03
Nov.99	8.25	68.06	16	256	132	.03
Dec.99	8.37	70.06	13	169	108.81	.00
Jan.00	8.5	72.25	11	121	93.5	.01
Feb.00	8.5	72.25	10	100	85	.01
Mar.00	8.73	76.21	10	100	87.3	.09
Apr.00	8.83	77.97	9	81	79.47	.16
May00	9	81	8	64	72	.32
Jun.00	9.24	85.38	7	49	64.68	.66
Jul.00	9.5	90.25	5	25	47.5	1.14
Aug.00	9.5	90.25	6	36	57	1.14
Sep.00	9.5	90.25	4	16	38	1.14
Oct.00	9.5	90.25	5	25	47.5	1.14
Nov.00	9.5	90.25	4	16	38	1.14
Dec.00	9.5	90.25	3	9	28.5	1.14
Jan.01	9.05	81.90	6	36	54.3	.38
Feb.01	8.5	72.25	8	64	68	.01
Mar.01	8.32	69.22	9	81	74.88	.01
Apr.01	7.8	60.84	14	196	109.2	.40
May01	7.24	52.42	16	256	115.84	1.42
Jun.01	7	49	20	400	140	2.04
	$\sum x =$ 252.89	$\sum x^2 =$ 2147.7749	$\sum y =$ 379	$\sum y^2 =$ 5955	$\sum xy$ = 3069.69	$\sum (x - \bar{x})^2$ = 15.9964967

Step 4: Calculate the intercept (a) of the regression equation

The next step is to calculate the intercept (a) of the regression equation. The formula is shown below:

Intercept (a)

$$a = \frac{\sum y}{N} - \frac{b\left(\sum x\right)}{N}$$

$$a = \frac{379}{30} - \frac{-7.823817(252.89)}{30} = 78.585505$$

Placing the appropriate sum totals of columns from Table 2 and performing the calculation results in an intercept (a) of 78.585505. The intercept represents the value of the performance indicator (y) when the value of the extraneous variable (x) is zero.

Step 5: Calculate the predicted value of the performance indicator (y') given an extraneous variable value (x).

Having calculated the slope (b) and intercept (a), these values can now be placed into the regression formula. However, before calculating the predicted value of the performance indicator (y') for the month of July 2001, the interest rate (x) for that month must be determined.

Obviously, then, the calculation of the predicted value can not occur until the actual interest rate for the month of July 2001 is known. Table 3 shows the interest rates (x) for July 2001 (the first month after training), as well as the interest rates for each month through April of 2002. The formula and subsequent calculation of the predicted value (y') of the performance indicator follow the table.

Table 3

Interest Rates for July 2001 to April 2002

Month	Interest Rate (in percent) (X)
Jul.01	7.0
Aug.01	7.0
Sep.01	7.25
Oct.01	7.25
Nov.01	7.5
Dec.01	7.67
Jan.02	7.67
Feb.02	7.67
Mar.02	8.0
Apr.02	8.05
May02	8.25
Jun.02	8.25

Regression Equation

$y' = a + b(x)$

$y' = 78.585505 + -7.823817(7.0)$

$y' = 23.819$

As shown above, the values of the intercept (a) and slope (b) have been placed into the regression equation. As can be seen in Table 3, the interest rate for July 2001 was 7.0. This value (x) was also placed into the regression equation resulting in a calculated value of y' being 23.819. The interpretation of this value is as follows: Given an interest rate of 7.0, the regression equation (based on 30 months of historical interest rates and loan data) predicts that the amount of loans sold by the 15 loan officers in July 2001 would have been $23,819 (23.819 x $1000; reflects figures in thousands) if they had not participated in the training.

At this point, the predicted value of $23,819 could be compared to the actual amount of loans that the newly-trained loan officers sold. However, making such a comparison at this point would be inappropriate. Suppose the actual amount of loans sold in July 2001 was $32,000, representing an increase over the predicted amount by $8,181 ($32,000 - $23,819). Is it time to celebrate and call the training a huge success? Not quite yet. What remains unanswered is the accuracy of the prediction. In other words, how much faith can be placed in the value derived from the regression equation? It must be remembered that y' is a "predicted" value. . . i.e., an estimate. So, let's continue.

Step 6: Calculate the Standard Error of Estimate (S_{ee})

The standard error of estimate provides an indication of the degree to which the actual values of y for the 30 months prior to training differ, on average, from those values that would have been predicted for the same months. The derived value of the standard error of estimate is somewhat difficult to interpret in terms of a 'good' value and a 'bad' value. However, the

standard error of estimate is necessary for calculations of values more easily interpreted. Those calculations will be discussed in the next several steps. The formula and calculations for the standard error of estimate are shown below.

Standard Error of Estimate (S_{ee})

$$S_{EE} = \sqrt{\frac{\sum y^2 - a\left(\sum y\right) - b\left(\sum xy\right)}{N-2}}$$

$$S_{EE} = \sqrt{\frac{5955 - 78.585505(379) - -7.823817(3069.69)}{30-2}} = 2.58973$$

Step 7: Calculate the Standard Error of Forecast (S_{ef})

Using the standard error of estimate value that was calculated in the previous step, the standard error of forecast can be derived. Stating it simply, the standard error of forecast indicates the degree to which predicted values of y (y') can be expected to vary from the actual values of y. Like the standard error of estimate, the derived value of the standard error of forecast is somewhat difficult to interpret in terms of a 'good' value and a 'bad' value. However, the standard error of forecast is necessary for calculation of a confidence interval, which is easily interpreted. The formula and calculations for the standard error of forecast are shown below.

Standard Error of Forecast (S_{ef})

$$S_{ef} = S_{ee}\left(\sqrt{1+\frac{1}{n}+\frac{\left(x_f-\bar{x}\right)^2}{\sum\left(x-\bar{x}\right)^2}}\right)$$

$$S_{ef} = 2.58973\left(\sqrt{1+\frac{1}{30}+\frac{(7.0-8.4297)^2}{15.9964967}}\right) = 2.7901629$$

Note in the formula $\left(x_f-\bar{x}\right)^2$; x_f is the actual interest rate for July 2001 used to predict the amount of loans sold (y') for that month and $\bar{x}$ is the mean of the interest rates (x) for the 30 months prior to training. Thus, the value that has been derived for the standard error of forecast (2.7901629) is the standard error of forecast for only the predicted value in July 2001.

To calculate the standard error of forecast for August 2001, and the remaining months after training, a predicted value of y (y') must first be calculated. In this calculation, the same slope (b) and intercept (a) values that were used to predict y' for July 2001 would be used again. However, the corresponding interest rate (x) for the month being predicted must be used in the regression equation. Each newly-predicted value, then, will serve as x_f in the calculation of the standard error of forecast for each month after training.

Step 8: Calculate a confidence interval around the predicted value

Once the standard error of forecast has been calculated, a confidence interval around the predicted value of y (y') can be constructed. The confidence interval provides a simple means of interpreting the accuracy of the predicted value. Specifically, the confidence interval represents a range of values that actual value of y would be expected to fall within.

As shown below, the calculations are simple. However, a decision must be made at this point whether to report a 68% confidence interval, a 95% confidence interval, or both. A 68% confidence interval has an expected probability of 68% that the actual value of y will fall within the interval, while a 95% confidence interval has an expected probability of 95% that the actual value of y will fall within the interval.

68% Confidence Interval

$CI_{68} = y' \pm S_{ef}$

$CI_{68} = 23.819 \pm 2.7901629 = 21.029$ to 26.609

95% Confidence Interval

$CI_{95} = y' \pm 1.96(S_{ef})$

$CI_{95} = 23.819 \pm 1.96(2.7901629) = 18.350$ to 29.287

We have already determined that, given the July 2001 interest rate of 7.0, the predicted amount of loans sold among the 15 loan officers in July would be $23,819 if they had not

participated in training. Thus, both 68% and 95% confidence intervals were constructed around this predicted value (23.819).

To calculate the 68% confidence interval, the standard error of forecast (S_{ef}) is added to the predicted value (y') to derive the upper range value of the interval, producing a value of 26.609 or $26,609. Next, the standard error of forecast (S_{ef}) is subtracted from the predicted value (y') to derive the lower range value of the interval, producing a value of 21.029 or $21,029. The interpretation of these figures is as follows: There is a 68% probability that the actual dollar amount of loans sold (y) for July 2001 will fall between $21,029 and $26,609.

A similar approach is taken to calculate the 95% confidence interval. However, you should note from the formula above that the standard error of forecast (S_{ef}) is first multiplied by the z value of 1.96. This value is a constant and should always be used in the calculation of a 95% confidence interval when N is equal to or greater than 30. From the calculations, a confidence interval of $18,350 to $29,287 is produced. The interpretation of these figures is as follows: There is a 95% probability that the actual dollar amount of loans sold (y) for July 2001 will fall between $18,350 and $29,287.

You should notice that this confidence interval is somewhat larger than the interval at the 68% confidence level. The tradeoff for greater confidence in your prediction is a wider range of possible actual values of y. Thus, judgement should be used regarding which confidence interval to report.

Step 9: Calculate the monetary benefits of the training program

To this point, the predicted value of y (y') and the confidence intervals around the predicted value have been calculated for July 2001. Table 4 shows the predicted values of y (y') and their respective 68% confidence intervals for the 11 additional months that followed training.

Also included in the table are actual interest rates for each month (x), the actual dollar amounts of loans sold by the 15 loan officers per month (y), and difference between the actual loan amounts sold and the upper limit of the confidence interval.

Table 4

Post-training Data

Month	Interest Rate (in percent) (X)	Total Loans Sold (in thousands) (Y)	Predicted Y	68% Confidence Interval	Actual Y After Training minus Upper Limit of Confidence Interval
Jul.01	7.0	32	23.819	21.029 to 26.609	5.391
Aug.01	7.0	34	23.819	21.029 to 26.609	7.391
Sep.01	7.25	30	21.863	19.122 to 24.603	5.396
Oct.01	7.25	29	21.863	19.122 to 24.603	4.396
Nov.01	7.5	28	19.907	17.207 to 22.607	5.393
Dec.01	7.67	26	18.577	15.899 to 21.254	4.745
Jan.02	7.67	27	18.577	15.899 to 21.254	5.745
Feb.02	7.67	29	18.577	15.899 to 21.254	7.745
Mar.02	8.0	25	15.995	13.348 to 18.642	6.358
Apr.02	8.05	22	15.604	12.960 to 18.247	3.753
May02	8.25	21	14.039	11.404 to 16.674	4.326
Jun.02	8.25	22	14.039	11.404 to 16.674	5.326
					Total = 65.965

As shown in the table, during the month of July 2001, the 15 loan officers sold $32,000 in loans. Based on historical data of interest rates and loans sold for 30 months prior to training, the regression equation that was calculated predicted, with the actual interest rate of 7.0, that the 15 loan officers would have sold $23,819 in loans. However, due to inherent error in this prediction, there is a 68% probability that actual loans sold for July 2001 would range between $21,029 and $26,609. Thus, to provide a more precise interpretation of the monetary benefit of

the training the actual loans sold by the loan officers should be compared to the upper limit of the confidence interval instead of the predicted value.

In this example, the difference between the actual loans sold in July 2001 and the upper limit of the confidence interval is $5,391. We can reasonably interpret this as follows: There is a 68% probability that the $5,391 in additional loan sales beyond the upper limit of the confidence interval is not the result of a change in the interest rate. In essence, the influence of the interest rate as an extraneous variable has been controlled and accounted for in the analysis.

Returning to Table 4, the difference between the actual loans sold and the corresponding upper limits of the confidence interval for each of the remaining 11 months are provided. During the 12 months following training, the total dollar value of loans sold beyond the confidence intervals is $65,965. Similar to the interpretation of the $5,391 for July 2001, this figure would be interpreted as follows: There is a 68% probability that the $65,965 in additional loan sales beyond the upper limits of the confidence intervals over the 12 month period following training is not the result of a change in the interest rates.

Step 10: Calculate Return on Investment

The final step in the process is to calculate the return on investment in the training program. The formula for the calculation is relatively straightforward.

Return on Investment

$$ROI = \frac{Program Benefits - Program Cost}{Program Cost} \times 100$$

$$ROI = \frac{\$5,391 - \$15,000}{\$15,000} \times 100 = -64.06\% \text{ for July 2001}$$

Recall from the introduction of this case example that the cost of the sales training program was $15,000. Further, $5,391 in loan sales above the upper limit of the confidence interval was generated for the month of July 2001. Placing these figures into the formula above results in a return on investment of -64.06%. Certainly, this is not good news. A negative return on investment never is good news.

It must be remembered, however, that the ROI figure was calculated using only the program benefits achieved after one month following training. As is often the case, it sometimes takes time for training to pay back in monetary benefits. In this example, a ROI is realized after only 3 months following the training, as shown below.

$$ROI = \frac{\$18{,}178 - \$15{,}000}{\$15{,}000} \times 100 = 21.19\% \text{ for July, August and September combined}$$

Thus, the monetary benefits of the training for the months of July, August and September of 2001 combined produced a ROI of 21.19%. In other words, for every dollar that was invested in the training program $0.21 was returned beyond the cost of the training.

Conclusion

The obvious question becomes whether the additional $5,391 in July 2001 was the result of the training, or something else. There is always a potential that other extraneous variables not included in the analysis may explain, at least partially, the improvement in the performance indicator. If it is believed that this is the case, then a multiple-regression approach in which

additional extraneous variables are added to the regression equation to predict y is warranted. However, such an approach is somewhat more complex.

On the other hand, if the evaluator provides additional evidence from Levels 1-3 evaluations that: 1) The training participants reacted favorably to the training and indicated how they planned to apply what was learned back on the job; 2) The training participants learned was they were supposed to learn and could apply skills during a training simulation, and; 3) The training participants were observed applying the knowledge and skills on the job, then there is strong evidence supporting the contention that any increases in the performance indicator beyond the upper limit of the confidence interval resulted from training.

References

Makridakis, S. & Wheelwright, S. C. (1989). *Forecasting methods for management* (5th ed.). New York, NY: John Wiley & Sons

Phillips, J. J. (1997). *Handbook of training evaluation and measurement methods* (3rd ed.). Houston, TX: Gulf Publishing Company

About the author:

C. Keith Waugh, Ph.D. is an assistant professor in the Department of Workforce Education and Development at Southern Illinois University - Carbondale. Dr. Waugh teaches graduate and undergraduate courses in training assessment and evaluation, and has been an evaluation consultant to numerous organizations.
Email: ckwaugh@siu.edu

CE083166

ACTE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Isolating the Effects of Training using Simple Regression Analysis: An Example of the Procedure	
Author(s): C. Keith Waugh	
Corporate Source:	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title: C. Keith Waugh, Asst. Professor	
Organization/Address: Southern Illinois Univ. at Carbondale	Telephone: 618-453-4868	FAX: 618-453-1909
	E-Mail Address: CKWAUGH@siu.edu	Date: 3/28/2002

212 Pulliam Hall
Carbondale, IL 62901-4605

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

Acquisitions Coordinator
ERIC Clearinghouse on Adult, Career, and Vocational Education
Center on Education and Training for Employment
1900 Kenny Road
Columbus, OH 43210-1090

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to: